Length Frequency

Alex J. Benecke

January 26, 2018

Data Preparation for Length-Frequency Analysis 2013 - 2017

I will now compare the length frequency distribution for largemouth bass obtained in the nearshore electrofishing survey during 2013 - 2017.

```
# lmb <-
# read.csv('Data/Clean-Data/2012-2017_nearshore-survey-largemouth-bass_CLEAN.csv')
# %>%
lmb <- read.csv("Data/Clean-Data/lmb-cpe.csv") %>% arrange(Year, FID, Length) %>%
    filter(Length >= 200)

lmb$fyr <- as.factor(lmb$fyr)
```

```
unique(lmb$Year) ### See that there is no 2013
```

```
## [1] 2013 2014 2015 2016
```

Lets create a new variable for 20 mm length bins.

`{ lcat20 2013 - 2017, tidy=TRUE} lmb %<>% mutate(lcat20=lencat(Length,w=20))`

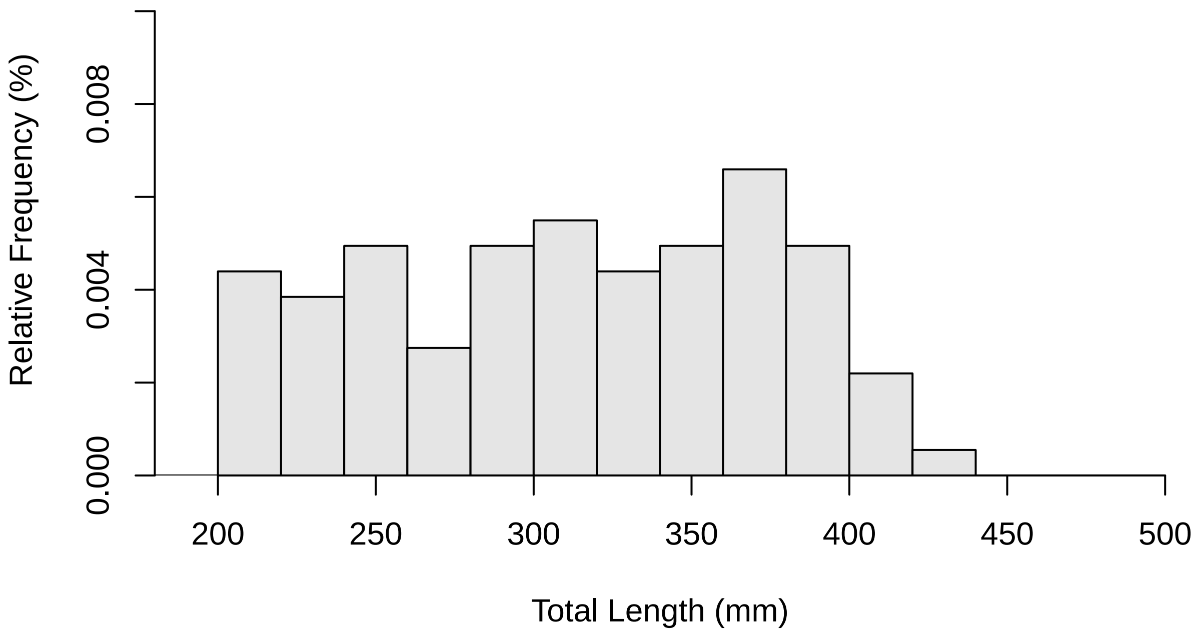
Now I want to separate out the years. I will throw out the year 2012 because samples from this years were not collected using the same procedures as in subsequent years. Only large LMB from 2012 had length weigh data.

```
# lmb.12 <- filter(lmb,Year==2012) 1-8-2018#write.csv(lmb.12, file =
# 'Data/Clean-Data/minor-data/lmb.12.csv')
lmb.13 <- filter(lmb, Year == 2013)
# 1-8-2018#write.csv(lmb.13, file = 'Data/Clean-Data/minor-data/lmb.13.csv')
lmb.14 <- filter(lmb, Year == 2014)
# 1-8-2018#write.csv(lmb.14, file = 'Data/Clean-Data/minor-data/lmb.14.csv')
lmb.15 <- filter(lmb, Year == 2015)
# 1-8-2018#write.csv(lmb.15, file = 'Data/Clean-Data/minor-data/lmb.15.csv')
lmb.16 <- filter(lmb, Year == 2016)
# 1-8-2018#write.csv(lmb.16, file = 'Data/Clean-Data/minor-data/lmb.16.csv')
# lmb.17 <- filter(lmb,Year==2017) 1-8-2018#write.csv(lmb.17, file =
# 'Data/Clean-Data/minor-data/lmb.17.csv')
```

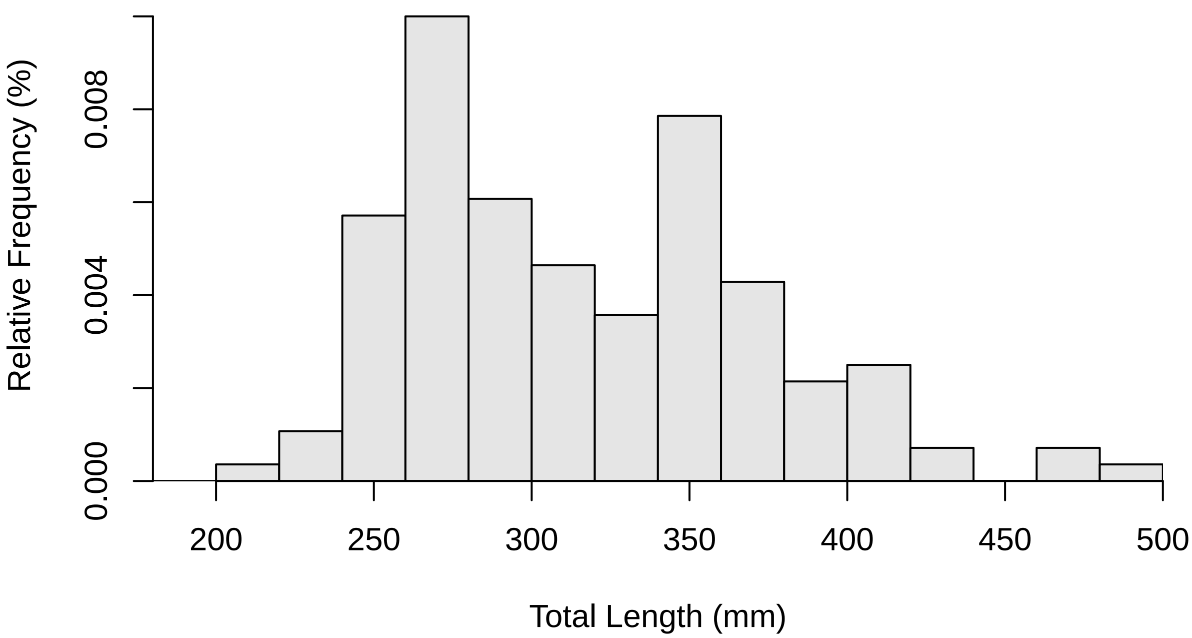
Length Frequency Distribution

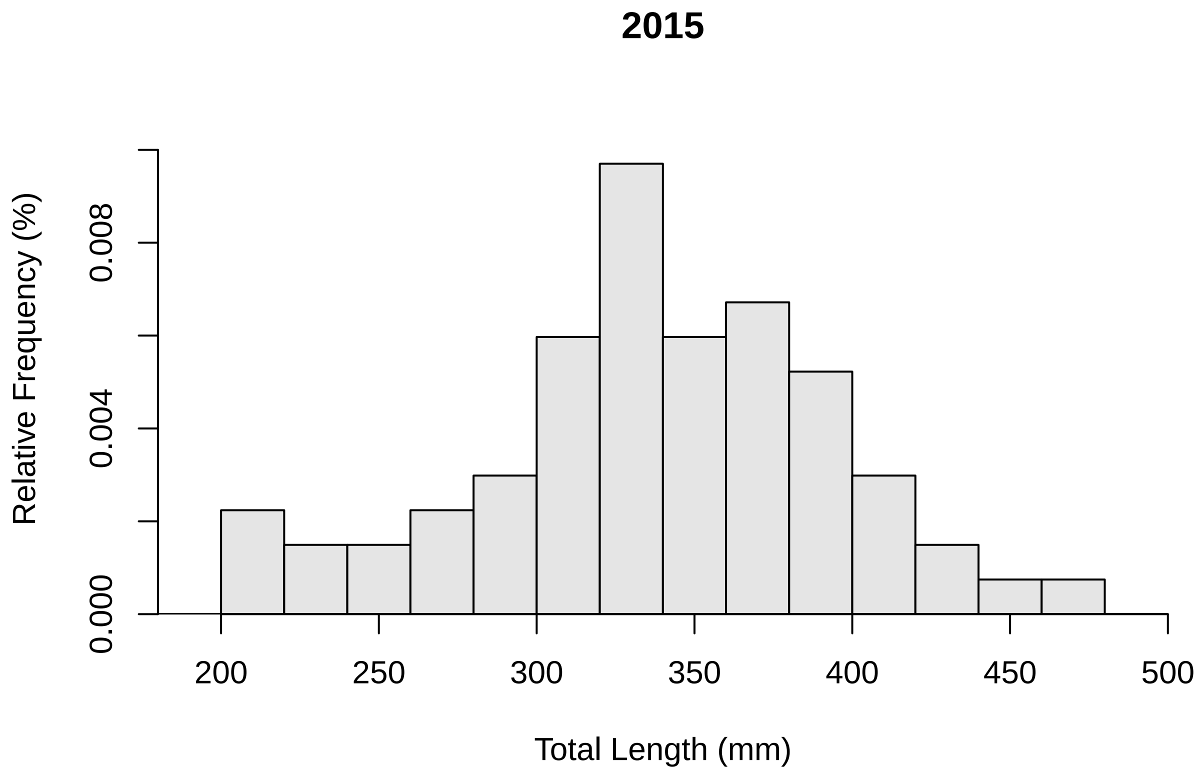
Lets view a quick histogram of the frequency of fish in each length bin.

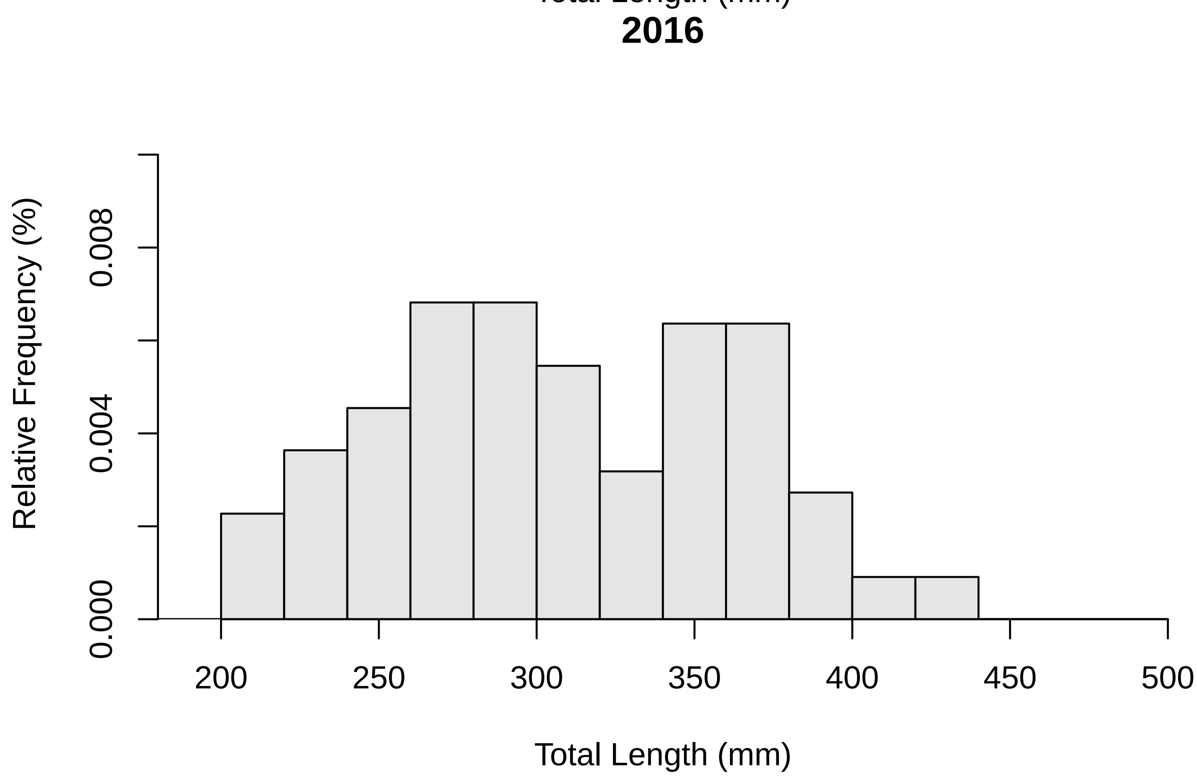
2013



2014







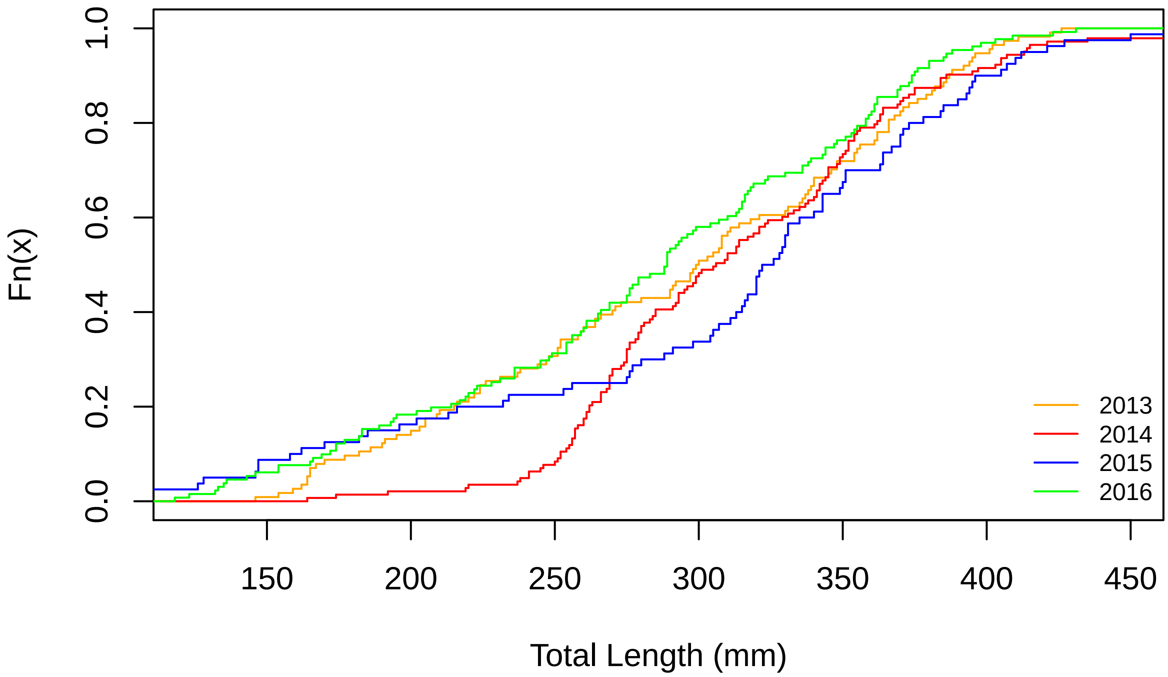
I can tell that small fish (<100-250mm) are definitely not being sampled consistently. **2014** and **2016** look more or less as I would expect with a stable largemouth bass population (2014 more so than 2016) with a peak and more or less gradual decline in the number of fish with increasing length. The peak for **2015** appears to be shifted right (320mm) more so than for 2014 and 2016 (260mm). The year **2013** Appears to be more or less flat but it's highest peak is at (360mm) before declining. As for **2017** I think this represents inadequate sampling I see a peak with much smaller fish (100mm) than other years and some length bins with only 1 fish. the peak for 2017 is around 340mm.

I would have to say from these graphs that the length frequency distribution of largemouth bass is unstable. However, I am tempted to think that there may be some aspect of sampling bias playing a role in this. **2017** was not sampled fully or by the same gear or crew (University of Toledo and old shock boat) due to staff shortages so I think I may throw this year out (missing 4 annual sites but has 2 rotating sites). **2013** is also missing several sites (6, 12, 19) although one site is NA so this could be one of those three. However, **2013** does seem to have a good sample size (n = 114) also site 19 was only sampled in 2017 and site 6 never seemed to provide alot of largemouth bass. So I will keep the year 2013 in at least for now. **2015** was also incompletely sampled missing two annual sotes (10 and 19) but with one rotating site for a total of nine sites. **2015** has a lower sample size than other years (n=80) *except for 2017*. However, since I dont't have any years in which all sites were sampled and the same number of sites were sampled in 2015 as in 2014 and 2016 I cannot throw out this year. Although, just looking at the data from 2015 it seems for whatever reason we have a significantly smaller proportion of smaller (stock and substock) fish.

Length Frequency 2013 - 2016

Cumulative Frequencies 2013 - 2016

Lets look at the empirical cumulative distribution function (ECDF). This is the proportion of fish less than each observed length. This should help me compare the length frequency distributions between years.



Compare Length Frequency Between Years 2013 - 2016

Kolmogorov-Smirnov Test

I will test wether the cumulative frequencies are different between years 2013 - 2016 using a Kolmogorov-Smirnov Two-tailed test. An example of this test that I used to help me teach myself how to do this can be found (https://ned.ipac.caltech.edu/level5/Wall2/Wal4_3.html). The formula I used to calculate the D value that would be significant at a given level of significance can be found (https://ned.ipac.caltech.edu/level5/Wall2/TableA_9.html).

Year	m	n	0.05	0.01	0.005	0.001
2013 - 2014	114	143	0.1707596	0.2046604	0.2172162	0.2448391
2013 - 2015	114	80	0.1983546	0.2377339	0.2523188	0.2844055
2013 - 2016	114	131	0.1741943	0.2087770	0.2215854	0.2497638
2014 - 2015	143	80	0.1898796	0.2275763	0.2415380	0.2722539
2014 - 2016	143	131	0.1644790	0.1971329	0.2092270	0.2358339
2015 - 2016	80	131	0.1929743	0.2312854	0.2454747	0.2766911

Compare Length Frequency Between Years 2013 - 2016

Kolmogorov-Smirnov Test

```
## [1] 0.23721016 0.19517544 0.09997322 0.19003497 0.24747771 0.24265267
```

```
## [1] "13-14" "13-15" "13-16" "14-15" "14-16" "15-16"
```

Summary of Results 2013 - 2016

The results of the Kolmogorov-Smirnov test above seem to suggest the largemouth bass population is not stable (*Or is it? I think there are just a few weird years probably sampling related*). The length frequency distribution **is significant different** between the years 2013 and 2014 (D = 0.24, p < 0.05), 2014 and 2016 (D = 0.25, p < 0.05), and 2015 and 2016 (D = 0.24, p < 0.05). There is **no significant difference** between the length frequency distributions for 2013 and 2015 (D = 0.20, p = 0.15), 2013 and 2016 (D = 0.10, p = 0.58), and 2014 and 2015 (D = 0.19, p = 0.15).

Year	D	D at p of 0.05	p-value	Relationship
2013 - 2014	0.23721	0.17076	0.008	S
2013 - 2015	0.19518	0.19835	0.148	NS
2013 - 2016	0.09997	0.17419	0.576	NS
2014 - 2015	0.19003	0.18988	0.148	NS
2014 - 2016	0.24748	0.16448	0.003	S
2015 - 2016	0.24265	0.19297	0.023	S

http://www.introspective-mode.org/kolmogorov-smirnov-goodness-of-fit/

Length Frequency 2013 - 2016 Stock Length (>= 200mm)

What is I remove small fish (<200mm) which I am not sure are fully sampled by my gear. Will 2014 still be different from 2013 and 2016? will 2015?

Year	n Sub-Stock	n Stock	n Quality	n Preferred
2013	16	41	41	16
2014	3	65	57	18
2015	13	14	38	15
2016	24	52	44	11

2014 has a lot of *stock* length individuals ($n_{stock} = 65$) compared to **2013** ($n_{stock} = 41$), **2015** ($n_{stock} = 14$), and **2016** ($n_{stock} = 52$) but far fewer *substock* individuals ($n_{sun-stock} = 3$), **2013** ($n_{sun-stock} = 16$), **2015** ($n_{sun-stock} = 13$), **2016** ($n_{sun-stock} = 24$). **2014** also seems to have a bit more but comperable amount of *quality* and *preferred* length individuals (57,18 respectivley) than **2013** (41,16), **2015** (38,15), and **2016** (44,11). So, lets explore wether the differences are in the smaller sub stock length not fully sampled by our gear.

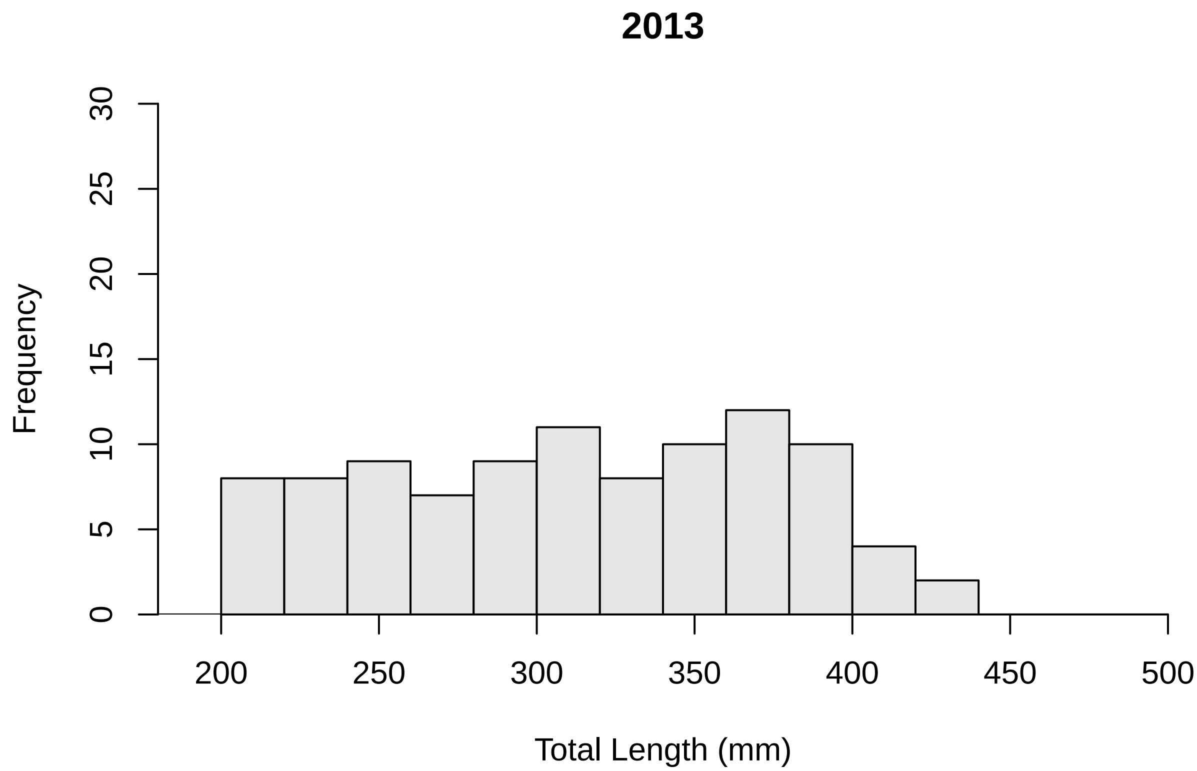
```
lmb.s <- read.csv("Data/Clean-Data/2012-2016_nearshore-survey-largemouth-bass_Stock_CLEAN.csv") %>%
    arrange(Year, FID, Length)
lmb.s$fyr <- as.factor(lmb.s$fyr)
```

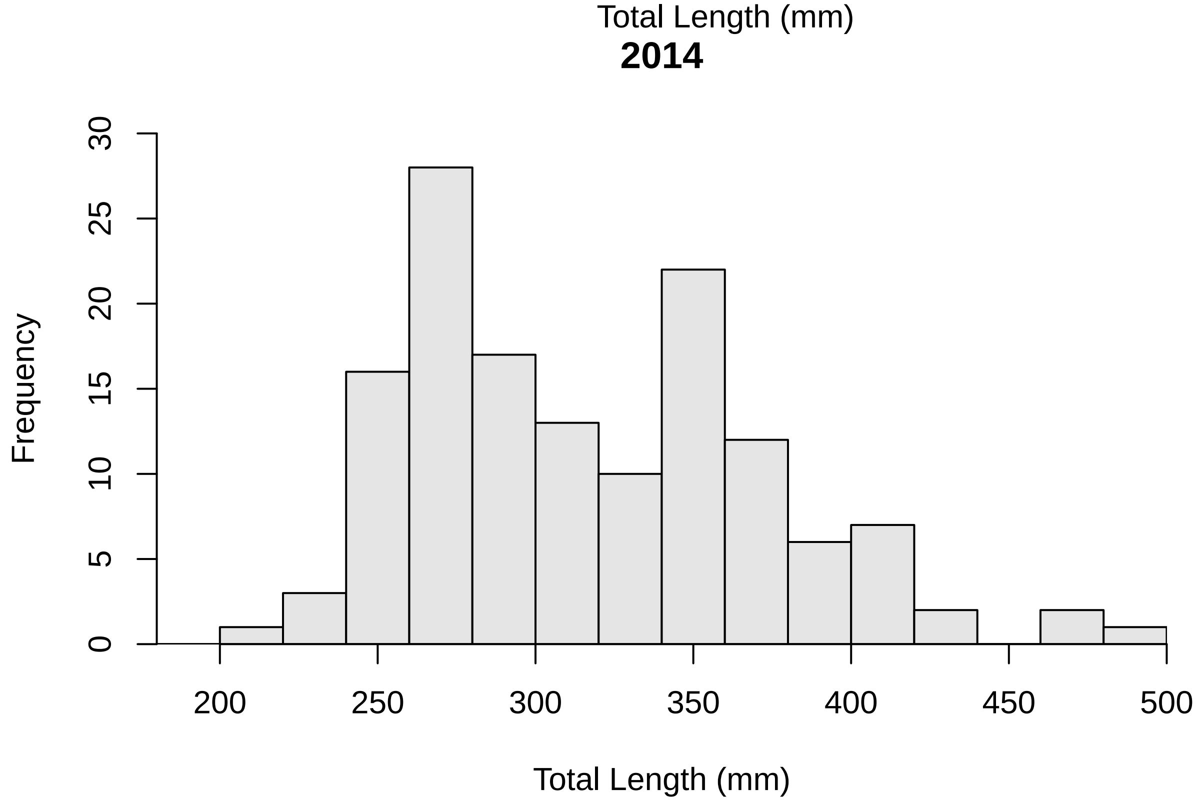
Year	m	n	0.05
2013 - 2014	98	140	0.1317078
2013 - 2015	98	67	0.1585227
2013 - 2016	98	107	0.1398209
2014 - 2015	140	67	0.1485538
2014 - 2016	140	107	0.1284081
2015 - 2016	67	107	0.1557921

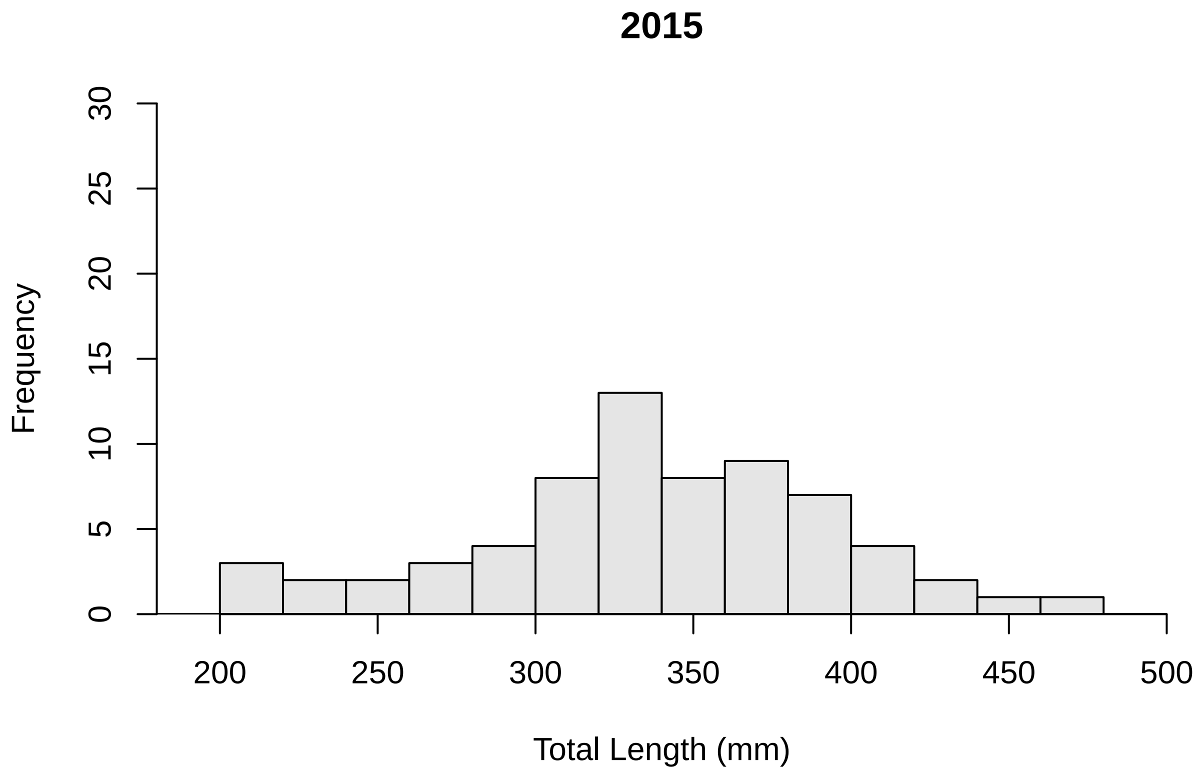
Length Frequency Distribution of Stock Length Fish During 2013 - 2016

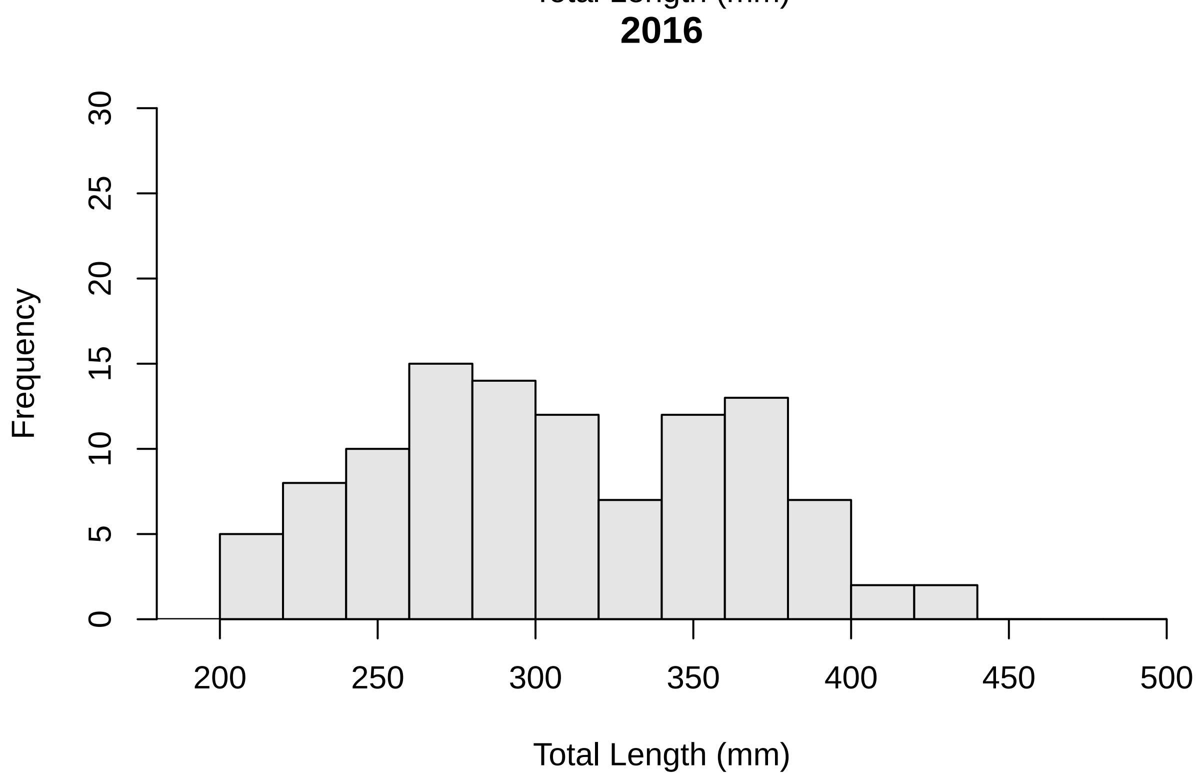
Length Frequency Distribution of Stock only fish 2013 - 2016

Lets view a quick histogram of the frequency of fish in each length bin.



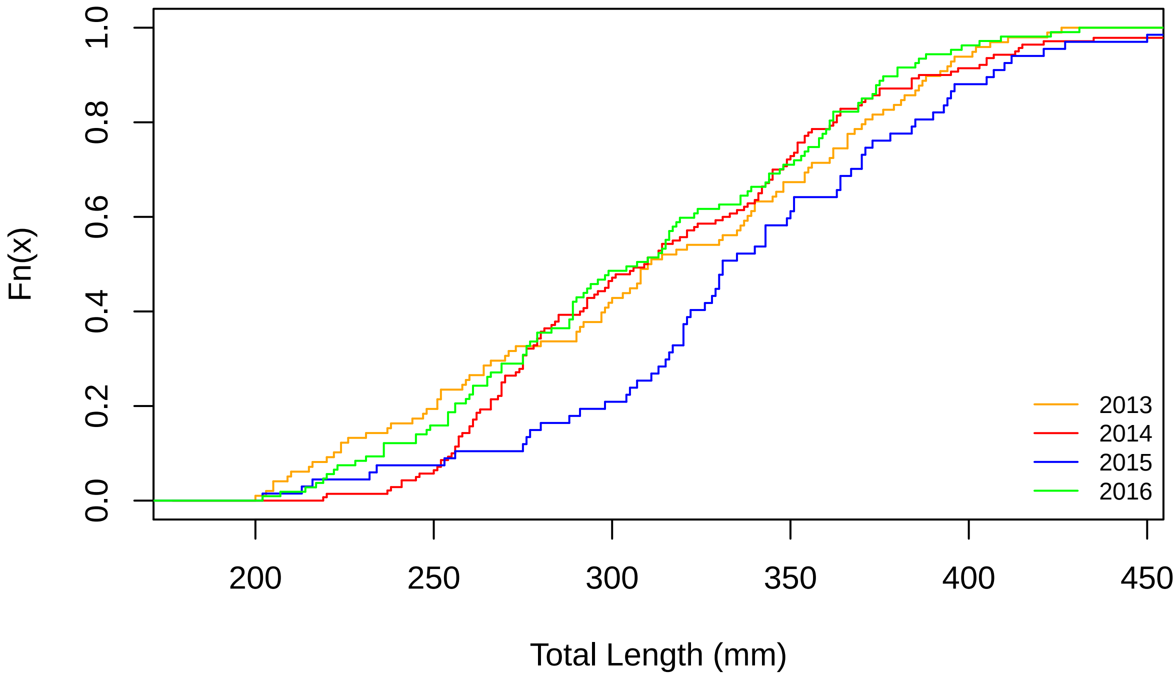



2015
Frequency
0
5
10
15
20
25
30
200
250
300
350
400
450
500
Total Length (mm)


2016
Frequency
0
5
10
15
20
25
30
200
250
300
350
400
450
500
Total Length (mm)

Cumulative Frequencies of Stock Length Fish Only During 2013 - 2016

Lets look at the empirical cumulative distribution function (ECDF). This is the proportion of fish less than each observed length. This should help me compare the length frequency distributions between years.



Compare Length Frequency Between Years 2013 - 2016 Stock Length Fish Only

Kolmogorov-Smirnov Test Of Only Stock Length Fish

```
## [1] 0.1489796 0.2462687 0.0897387 0.2696162 0.1072096 0.2770261
```

```
## [1] "13-14" "13-15" "13-16" "14-15" "14-16" "15-16"
```

Summary of Results 2013 - 2016 Stock Length Fish Only

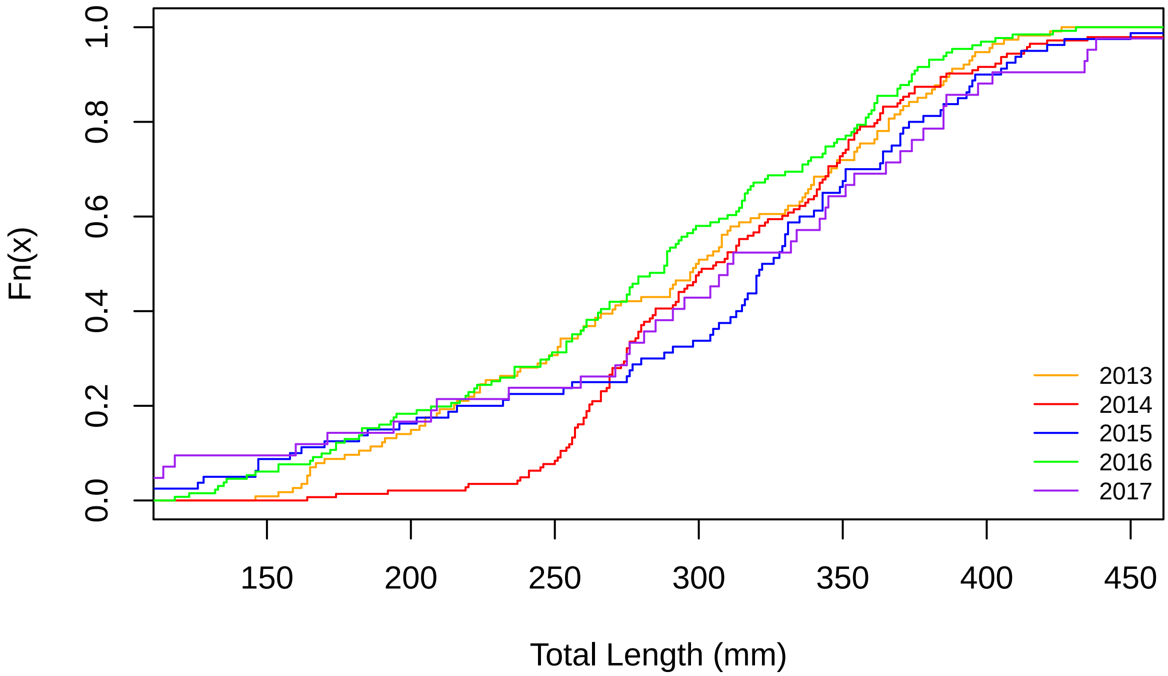
The length frequency distribution **is significant different** between the years 2014 and 2015 (D = 0.27, $p < 0.05$), and 2015 and 2016 (D = 0.28, $p < 0.05$). There is **no significant difference** between the length frequency distributions for 2013 and 2014 (D = 0.15, p = 0.46), 2013 and 2015 (D = 0.25, P = 0.06), 2013 and 2016 (D = 0.09, P = 0.98), and 2014 and 2016 (D = 0.11, P = 0.98).

Year	D	D at alpha of 0.05	p-value	Relationship
2013 - 2014	0.14898	0.17912	0.46	NS
2013 - 2015	0.24627	0.21559	0.06	~S
2013 - 2016	0.08974	0.19016	0.98	NS
2014 - 2015	0.26962	0.20203	0.02	S
2014 - 2016	0.10721	0.17464	0.98	NS
2015 - 2016	0.27703	0.21188	0.02	S

2013 - 2017 Length Frequency

Cumulative Frequencies 2013 - 2017

Lets look at the empirical cumulative distribution function (ECDF). This is the proportion of fish less than each observed length. This should help me compare the length frequency distributions between years.



Summary of Results 2013 - 2017

The results of the Kolmogorov-Smirnov test above seem to suggest the largemouth bass population is not stable (*Or is it? I think there are just a few weird years probably sampling related*). The length frequency distribution **is significant different** between the years 2013 and 2014 (D = 0.24, P = 0.014), 2014 and 2016 (D = 0.25, P < 0.005), and 2015 and 2016 (D = 0.24, P = 0.046). There is **no significant differnce** between the length frequency distributions for 2013 and 2015 (D = 0.20, P = 0.344), 2013 and 2016 (D = 0.10, P = 1), and 2014 and 2015 (D = 0.19, P = 0.344). The length frequency distribution for the year 2017 was *not significantly different* between any years.

Note: *Adding in the Year 2017 Significantly Altered the Adjusted P-Values*

Year	D	p-value	Relationship
2013 - 2014	0.24	0.014	D
2013 - 2015	0.20	0.344	S
2013 - 2016	0.10	1.000	S
2013 - 2017	0.14	1.000	S
2014 - 2015	0.19	0.344	S
2014 - 2016	0.25	0.005	D
2014 - 2017	0.20	0.686	S
2015 - 2016	0.24	0.046	D
2015 - 2017	0.14	1.000	S
2016 - 2017	0.17	1.000	S